EXPLORING THE POTENTIAL OF OPEN SOURCE DATA TO GENERATE CONGESTION AND EMISSION TRENDS IN DEVELOPING CITIES

T. Moyo[a, *], A. Kibangou[b], W. Musakwa[c]

[a] Dept. of Quality and Operations Management, University of Johannesburg, Doornfontein Campus, 0184 Johannesburg, South Africa
[b] Université Grenoble Alpes, Domaine Universitaire - BP46, 38402 Saint Martin d'Hères, France
[c] Future earth and Ecosystem Services Research Group, University of Johannesburg, Doornfontein, 0184 Johannesburg, South Africa

KEY WORDS: Emission, Congestion, COVID-19, TomTom data, Johannesburg

ABSTRACT:

The growth in Intelligent Transportation Systems (ITS) has enhanced the way mobility in contemporary cities is managed. Given the growth in availability of traffic data that contains location-aware datasets, congestion and pollution indexes can be developed. Metropolitan cities such Johannesburg due to their economic activities, attract commuters into the city on a daily basis seeking greener pastures. This has led to major freeways and roads experiencing high levels of congestion. In 2020, due to a global pandemic of an outbreak of Corona Virus (COVID-19), the national government declared a national shutdown with only essential traffic being allowed to operate. Given the scenario of the national lock-down this allows for the statistical analysis of the impact of essential traffic on the overall transportation system. Consequently the aim of the paper was to explore the congestion and C02 emission impact of essential traffic for the City of Johannesburg. Using an exploratory approach, we monitored and collected traffic congestion data from the Tomtom traffic index for the metropolitan city of Johannesburg, South Africa. Using a mathematic model, we develop a relationship between congestion and pollution to visualise the variations in pollution and congestion levels during varies scenarios. We demonstrate this by comparing datasets for variations in congestion levels in two epochs, viz the period without movement restrictions and the period whereby movement is restricted. The results reveal essential traffic on the congestion index to be below 22 percent for both weekends and weekdays. A scenario common only during weekends in 2019. Whilst for the emission index, C02 levels are approximately less than 45 percent throughout the week. The paper concludes the investment into mining and analysing traffic data has a significantly role for future mobility planning in both the developed and developing world and, more generally, improving the quality of commuting trips in the city.

1. INTRODUCTION

Operation efficiency and environmental preservation have over the years become priorities in mobility planning. In 2020 a global pandemic of the coronavirus (COVID-19) led to many developing and developed countries closing down their economic operations. The virus spreads mainly through respiratory droplets which are transferred to another individual through close personal contact. To limit the spread of the virus and control transmission rates to allow heath practitioners to improve health care systems national governments have restricted movement rights of citizens. Due to this limitation only essential traffic is permitted during the national lock-down. From an operational perspectives this allows traffic authorities to assess the statistical impact of essential traffic on the overall transportation system. Operational efficiency can be evaluated in terms of the congestion index and environmental preservation in terms of the emission index.

To identify contributing factors leading to congestion on major highways scholars have outlined the road condition, traffic volume and speed as the foremost aspects leading to congestion (Polzin, 2017). Generally, traffic congestion can be defined as the additional travel time taken by vehicles when they traverse over a given spatial location when compared to the norm free-flowing travel time (Zhao and Hu, 2019). Galatioto et al., (2014) articulates the relationship between congestion and pollution levels in cities (Galatioto et al., 2014). Over the years various factors have been identified that led traffic congestion namely micro-level (relating to current conditions of roads) and macro-level (relating to demand for road usage) (Rao and Rao, 2012).

*Corresponding author- thembanijoel@gmail.com

Conversational congestion and emission indexes are undertaken using either macroscopic models, microscopic models or mesoscopic models. Such models rely on either observed vehicle engine, vehicle acceleration data or standard emission coefficients to monitor congestion or emission trends. In the age of open source data, almost every aspect of contemporary mobility research is impacted by the wealth of big data. Big data is generally expressed by the characteristics of the information, that is variability, volume, visualisation, velocity and veracity (Lorenzi et al., 2014). Traffic index data promises to assist transportation planners in both developed and developing countries with reliable information to inform mobility models which inform decision making. Perhaps the inco-operation of traffic data on congestion trends can be used in conjunction with emission co-efficient to develop an emission index for the city of Johannesburg. Consequently the aim of the paper was to explore the congestion and C02 emission impact of essential traffic for the City of Johannesburg.

2. LITERATURE REVIEW

2.1 Congestion measurement

Generally congestion is measured as a ratio of volume-to-capacity. In mobility studies the best practice appears to be using accessible data resources within an analysis framework that can eventually capture the benefits of improved or ideal data. An example of this practice comes from the input-output approach (Michalopoulos and Pisharody, 1981) (Liu et al., 2009). Vehicle queues at intersections are expressed as a function of traffic capacity and demand. Such a framework is normally utilised at traffic-controlled intersections to ensure a constant average of permissible congestion is allowed to ensure a balance between the traffic jammed flow and capacity flow (Michalopoulos and Pisharody, 1981). Other congestion estimates studies have sought to detect queue lengths in real-time using location or trajectory data of probe vehicles. (Bremmer et al., 2004) proposed using operational data collected from sensors placed along major highways to estimate traffic volumes and speeds. These sensors were used to gather vehicle count and travel time data between sensors. Findings from the analysis were to be able to distinguish between recurring congestion (moderately acceptable congestion triggered by routine traffic volumes operating in a typical setting) and nonrecurring congestion (unforeseen congestion triggered by unpredictable actions such as inclement weather or traffic accidents) at real-time. (Dias et al., 2009) proposed to measure real-time congestion one should consider:

$$C = \frac{V_f - V}{V_f} \tag{1}$$

where C represents Congestion, V_f free flow speed and V is actual travel time or speed of the vehicle. Compared to utilising one fixed travel speed, using the actual travel speed has the merit of a consistent representative of current congestion trends. To calculate actual travel speeds, detectors are used. These can be placed at pre-defined locations along roads or by monitoring the Global Positioning System (GPS) signal from vehicles. Mounting speed detectors along roads although expensive allows for city authorities to be able to collect ground true data on major roads. (Alexander and González, 2015) estimated travel patterns by extracting the average daily origin-destination trips from mobile phone records. Finding from the study reveal the modal shift to rideshare services led to has the potential to significantly reduce congestion on major roads, however this varies spatially and temporally, as it is subject to the distribution of trips and mode shares. The methods used in these studies have been effective in disseminating and sharing the traffic congestion trends for varies spatial locations, however, are limited to the specific spatial locations as setting up sensors along roads maybe hindered by privacy concerns. Based on such a situational assessment, additional sources of travel pattern data have been inco-operated by other studies and traffic agencies. The following simplified function ϵ congestion index is expressed as the ratio of extra travel time ΔT with respect to the free flow travel time T_f:

$$\epsilon = \frac{\Delta T}{T_f}. \tag{2}$$

Using a algebraic operations we then link the average speed V with the free flow speed V_f as

$$\frac{V_f}{V} = 1 + \epsilon \tag{3}$$

A notable example is Tomtom (a company that provides location technology and consumer technologies), which has developed a real-time congestion index based on such GPS data from vehicles and sensors along major roads. This presents a robust open source data resource which researchers and city authorities can utilise to analyse and visualise mobility trends over time. The growth in such open source platforms promises to create new opportunities, of congestion analysis, in this paper we hence utilise data from the traffic index as an input to unpack traffic trends for the city of Johannesburg.

2.2 Emission measurement

Contemporary research to mitigate problems emanating from the transportation systems has rapidly grown. Globally the dominant anthropogenic sources of air pollution have been emissions from vehicular transportation. Due to the high volumes of vehicles in urban areas a vast body of mobility research has been developed to measure and evaluate the pollutants exhaust emissions from vehicles (Boriboonsomsin et al., 2012)(Song et al., 2013). As the reduction and estimation of these emissions will led to a better quality of life in urban areas. In (Boriboonsomsin et al., 2012) the authors derive Energy/Emissions Operational Parameter Set (EOPS) from a large data set of speed profiles with each record with characteristics of time, geographic location and travelling speed. Using macroscopic traffic parameters the coefficients in Table of Emission are expressed as

$$\ln(E) = \sum_{k=0}^{4} \alpha_k V^k + bg \tag{4}$$

where g depends on the gradient of the road and is given in percentage. It can be rewritten according to the congestion index ϵ as:

$$\ln(E) = \sum_{k=0}^{4} \alpha_k \frac{V_f^k}{(1+\epsilon)^k} + bg \tag{5}$$

3. STUDY AREA

The Metropolitan City of Johannesburg is the economic hub of the wealthiest province, Gauteng in South Africa (see Figure 1). Due to these economic activities, many people traverse to this city on a daily basis seeking greener pastures. This has led to major freeways and roads experiencing high levels of congestion. In 2020, due to a global pandemic of an outbreak of Corona Virus (COVID-19), the national government declared a national shut down on the 23 of March 2020 which only came into effect on the 23 of March 2020 (Gilbert et al., 2020). As part of the shutdown, movement between provinces, metros and district areas was banned. To facilitate movement by essential workers public transport for essential trips was restricted, in the morning between 05:00 to 09:00 and evening between 16:00 to 20:00. On the 1st of April 2020 the passenger loading capacity on vehicles was revised from 50 percent to 70 percent (Nyabadza et al., 2020). Table 1 summarises the timeline of movement restrictions imposed in South Africa since the first reported COVID-19 case. In the paper we hence seek to visualise mobility trends that have occurred due to such restriction to travel in the city of Johannesburg.

Table 1: COVID-19 Movement Restriction Timeline in South Africa

Date	Event
March 5	First recorded COVID-19 case in South Africa
March 15	President declares a national state of disaster
March 16	Ports of entry and border ports into South Africa closed
March 18	Travel ban on foreign nationals from high-risk countries
March 26	National lock-down
April 1	Public transportation operations revised

4. METHODOLOGY

We use data collected from the Tomtom Traffic Index for the City of Johannesburg, South Africa for a one week period from

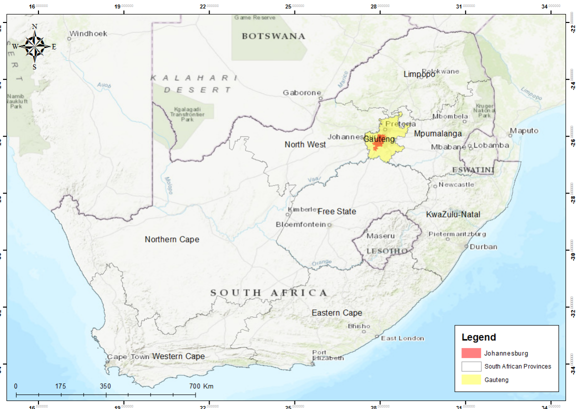

Figure 1: Study Area

14/04/2020 to 20/04/2020 and compare this data with the previous year in 2019 for the same time period 14/04/2019 to 20/04/2019.

For a given Energy/Emission Operational Parameter Set (EOPS), E, we denote by E_m its minimum value, V_m the corresponding average speed, and ϵ_m the congestion level corresponding to V_m. The pollution index is then defined as

$$\gamma(\epsilon, V_f) = \frac{E}{E_m} - 1. \tag{6}$$

Using (Boriboonsomsin et al., 2012) and (Song et al., 2013) to model Emission of light vehicles we derive EOPS as

$$E = \frac{a}{V} + \sum_{k=0}^{2} \beta_k V^k. \tag{7}$$

Parameters a and β_k depends on the type of the vehicle and the type of EOPS (Fuel consumption, HC, NOx, CO_2, and CO).

Regression coefficients for Light duty vehicles (Song et al., 2013)

CO_2

a=4780; β_0=111; β_1=-1.24; β_2=2.37×10^{-2}

We can therefore show that:

$$\frac{E}{E_m} = \frac{\Phi(\epsilon, V_f)}{\Phi(\epsilon_m, V_f)}, \tag{8}$$

where

$$\Phi(\epsilon, V_f) = \left(a(1+\epsilon) + \sum_{k=0}^{2} \frac{\beta_k V_f^{k+1}}{(1+\epsilon)^k} \right) \tag{9}$$

Given the city of Johannesburg road network which consists in urban arterial and highways we limit the free-flow speed to 50 Km/hr. For EOPS using coefficients for $CO2$ with a maximum free-flow speed 50km/hr we depict the pollution index increasing with the congestion index based for the sample data collected on Monday 20/04/2020 and 20/04/2019 see Fig **??** and **??**.

Free Flow $V_f = 50$ Km/h

$$\gamma(\epsilon, 50) = 0.0269 + 0.4752\epsilon - \frac{0.3082}{1+\epsilon} + \frac{0.2945}{(1+\epsilon)^2}$$

5. RESULTS

Figure 2 depicts the variations in congestion levels in two epochs, viz the period without movement restrictions and the period whereby movement is restricted. In 2019 according to the Tomtom Traffic Congestion Index for the city of Johannesburg Figure2 for weekdays congestion averaged between 28 percent to 69 percent during the morning peak hour 06:00 to 08:00 and 59 percent to 13 percent during the afternoon peak hour between 15:00 to 18:00. For weekends congestion averaged between 2 percent to 13 percent during the morning peak hour 06:00 to 08:00 and 8 percent to 12 percent during the afternoon peak hour between 15:00 to 18:00. This statistic is important in estimating the impact the national shutdown has had on congestion in the city. It is important to note that at the beginning of the epidemic, mobility in the country was considered uninterrupted even through concerns over the global impact of COVID-19 were present. The data from the Tomtom Traffic Congestion Index was then collected and visualised for the period from the 14/04/2020 to 20/04/2020 approximately 18 days after the declaration of a national shutdown.

Figure 3 depicts an abrupt change in congestion trends in the city. For weekdays congestion averaged between 8 percent to 12 percent during the morning peak hour 06:00 to 08:00 and 11 percent to 21 percent during the afternoon peak hour between 15:00 to 18:00. For weekends congestion averaged between 1 percent to 8 percent during the morning peak hour 06:00 to 08:00 and 1 percent to 13 percent during the afternoon peak hour between 15:00 to 18:00. The uniqueness of the trend in the 2020 data is that congestion levels are all below 22 percent for both weekdays and weekends. A scenario common only during weekends in 2019. The low levels of in congestion trends in 2020 between weekdays and weekends could be due to the restricted movement with only movement deemed essential being allowed. Also, as part of the lock-down restrictions most trips in the city can be described as short trips, that is between home and the local commercial shops, which generally do not result in high congestion.

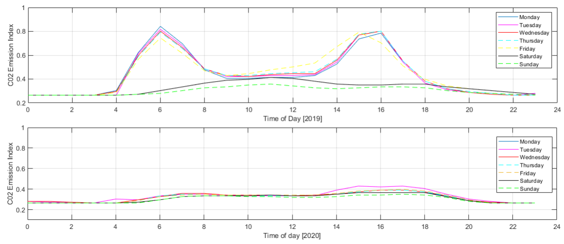

Figure 2: Comparison of CO_2 emission index of 2019 and 2020

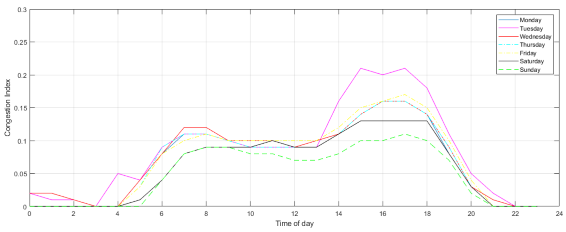

Figure 3: Time evolution of Congestion emission index

5.1 EOPS Applicability to the City of Johannesburg

We then consider the $CO2$ coefficients to assess the applicability of EOPS for the city of Johannesburg. The Emission Index is a useful tool for assessing emission rates and informing city planners on possible deterring measures. Figure 4 summarise the trends in CO2 emissions during the lock-down period in 2020. As noted in the previous section at 50km/hr the emission index increasing with the congestion index. This leads to a similar distribution trend. Figure 4 in 2019 for weekdays emission averaged between 81 percent to 48 percent during the morning peak hour 06:00 to 08:00 and 79 percent to 37 percent during the afternoon peak hour between 15:00 to 18:00. For weekends emission averaged between 21 percent to 39 percent during the morning peak hour 06:00 to 08:00 and 38 percent to 36 percent during the afternoon peak hour between 15:00 to 18:00. Given how most trips made during the weekend are for recreational purposes the general distribution of emission levels is even throughout the day.

For the data between 14/04/2020 to 20/04/2020 emission is approximately less than 45 percent through the week, this trend is similar to the congestion index (see Figure 5). For weekdays emissions averaged between 33 percent to 35 percent during the morning peak hour 06:00 to 08:00 and 34 percent to 43 percent during the afternoon peak hour between 15:00 to 18:00. For weekends congestion averaged between 29 percent to 33 percent during the morning peak hour 06:00 to 08:00 and 33 percent to 37 percent during the afternoon peak hour between 15:00 to 18:00.

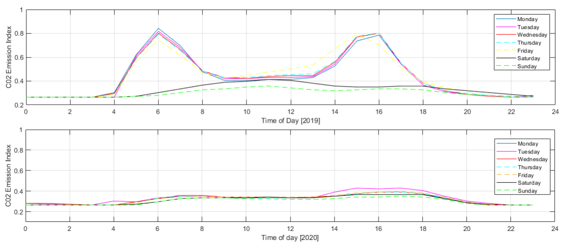

Figure 4: Comparison of CO_2 emission index of 2019 and 2020

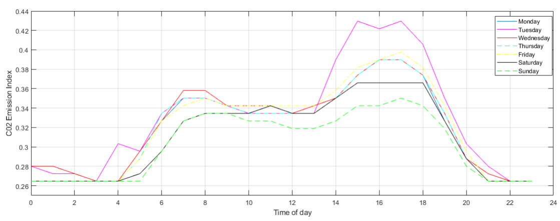

Figure 5: Time evolution of CO_2 emission index

To compare variations and similarities between the two trends in the two years 2019 and 2020 we use a cumulative distribution plot. For the weekend in 2019 approximately 90 percent is between 33 percent and 38 percent, while for weekdays is 70 percent. The contribution of the national lock-down to reducing emission can best be visualised by Figure 7 in comparison with the previous year. For all days of the week C02 emission is between 33 percent to 44 percent.

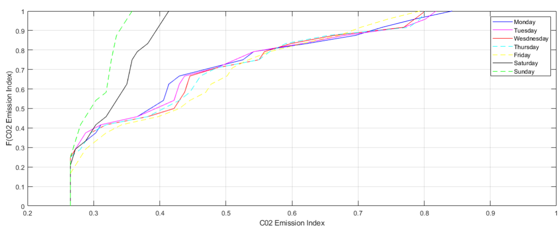

Figure 6: Cumulative distribution function of CO_2 emission index in 2019

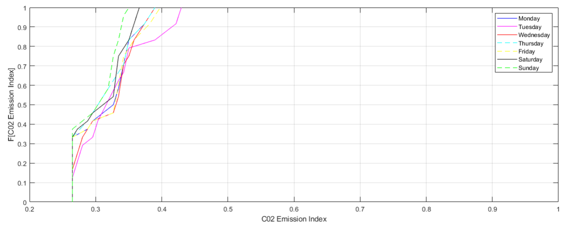

Figure 7: Cumulative distribution function of CO_2 emission index in 2020

6. CONCLUSION

The paper has presented a novel approach of utilising open source data as an input to inform mobility planning. The result reveal in 2020 the reduced levels of emission during the national lock-down reveal if mobility in the city is restricted, this will lead to reduced better quality of air. However, introducing a lock-down is not always an advisable option to reduce vehicular emissions as this would negatively impact economic activities. We however propose shifting commuting trips to public transportation or ride-sharing services, this would reduce the volume of traffic on highways whilst also reducing vehicular emissions. The investment in mining and analysing open source traffic data has a significantly role for future mobility planning in both the developed and developing world and, more generally, improving the quality of commuting trips in the city.

ACKNOWLEDGEMENTS

This work was funded by The National Research Foundation, South Africa (Grant No 110778) and National Research Foundation, South Africa (Grant no. 119288)

REFERENCES

Alexander, L. P. and González, M. C., 2015. Assessing the impact of real-time ridesharing on urban traffic using mobile phone data. Proc. UrbComp pp. 1–9.

Boriboonsomsin, K., Barth, M., Zhu, W. and Vu, A., 2012. Eco-routing navigation system based on multisource historical and real-time traffic information. IEEE Trans. on Intelligent transportation Systems 13(4), pp. 1694–1704.

Bremmer, D., Cotton, K. C., Cotey, D., Prestrud, C. E. and Westby, G., 2004. Measuring congestion: Learning from operational data. Transportation Research Record 1895(1), pp. 188–196.

Dias, C., Miska, M., Kuwahara, M. and Warita, H., 2009. Relationship between congestion and traffic accidents on expressways: an investigation with bayesian belief networks. In: Proceedings of 40th annual meeting of infrastructure planning (JSCE).

Galatioto, F., Bell, M. C. and Hill, G., 2014. Understanding the characteristics of the microenvironments in urban street canyons through analysis of pollution measured using a novel pervasive sensor array. Environmental monitoring and assessment 186(11), pp. 7443–7460.

Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boëlle, P.-Y., d'Ortenzio, E., Yazdanpanah, Y., Eholie, S. P., Altmann, M. et al., 2020. Preparedness and vulnerability of african countries against importations of covid-19: a modelling study. The Lancet 395(10227), pp. 871–877.

Liu, H. X., Wu, X., Ma, W. and Hu, H., 2009. Real-time queue length estimation for congested signalized intersections. Transportation research part C: emerging technologies 17(4), pp. 412–427.

Lorenzi, D., Vaidya, J., Shafiq, B., Chun, S., Vegesna, N., Alzamil, Z., Adam, N., Wainer, S. and Atluri, V., 2014. Utilizing social media to improve local government responsiveness. In: Proceedings of the 15th annual international conference on digital government research, pp. 236–244.

Michalopoulos, P. G. and Pisharody, V. B., 1981. Derivation of delays based on improved macroscopic traffic models. Transportation Research Part B: Methodological 15(5), pp. 299–317.

Nyabadza, F., Chirove, F., Chukwu, W. C. and Visaya, M. V., 2020. Modelling the potential impact of social distancing on the covid-19 epidemic in south africa. medRxiv.

Polzin, S., 2017. Vehicle miles traveled trends and implications for the us interstate highway system. Transportation Research Board, Washington, DC.

Rao, A. M. and Rao, K. R., 2012. Measuring urban traffic congestion-a review. International Journal for Traffic & Transport Engineering.

Song, Y.-Y., Yao, E.-J., Zuo, T. and Lang, Z.-F., 2013. Emissions and fuel consumption modeling for evaluating environmental effectiveness of ITS strategies. Discrete Dynamics in Nature and Society.

Zhao, P. and Hu, H., 2019. Geographical patterns of traffic congestion in growing megacities: Big data analytics from beijing. Cities 92, pp. 164–174.